

PLAINセンターニュース

Center for PLAnning and INformation Systems

仮想天文台 (Japanese Virtual Observatory) について

1. はじめに

様々な望遠鏡や観測装置で観測された宇宙の姿は、世界中の計算機の中に数値データ（これを観測された「数値宇宙」と呼ぶ）として記録されている。従って、これまでに観測された宇宙は計算機ネットワークを通して再観測を行うことができる。この数値宇宙を観測する望遠鏡を仮想望遠鏡 (Virtual Telescope) といい、数値宇宙の一部と仮想望遠鏡をもつ天文台を仮想天文台 (VO: Virtual Observatory) という。数値宇宙は、スローン・デジタル・スカイ・サーベイ (SDSS) のように専用望遠鏡による大規模なサーベイ観測が盛んに行われるようになり急速に拡大している。これら多波長のサーベイ観測データの比較研究によって、これまで見つけることが難しかった希な天体が発見できるようになり、データベース天文学という新たな天文学が誕生した。VO では観測波長に依存しない仮想望遠鏡によって容易に多波長の観測データの比較研究が行える。VO はデータベース天文学のための強力な道具であり、これによってデータベース天文学が発展すると期待している。

国立天文台では日本版 VO (JVO) の開発を行っており、仮想天文台とデータベース天文学については、天文月報 2002年 6,7 月号のデータベース天文学特集で紹介されている。これらの記事は JVO のホームページ (<http://jvo.nao.ac.jp/>) に掲載されているので参照されたい。ここでは、国立天文台データベース天文学推進室で昨年からの開発を進めている JVO プロトタイプについて紹介する。

2. JVO プロトタイプ

JVO の開発ではオブジェクト指向開発手法のスパイラル方式を採用している。実証したい基本機能のみを実装したプロトタイプを作成し、採用した技術の有効性や機能の評価を行い、機能を追加し実装を改良した次のプロトタイプを作成する。これを何度か繰り返し、最終的に運用システムを構築する。現在は、最初のプロトタイプの作成が終わり評価を行っているところである。

2.1 プロトタイプの目標

JVO の最初のプロトタイプを開発するにあたり、

- ・分散データベースを統一的に扱う方法を開発
- ・分散環境の基盤技術として GRID を採用

の2点を目標にした。概念設計にあたり、JVO を使って何をするのか実際の例 (Use Case) を数え上げる。例えば、「2つ以上のバンドで取られている領域からバンド/カラーごとの star count を求め銀河系のモデルと比較する」といったものである。この Use Case の実際の処理を机上でシミュレーションして必要な処理を洗い出す。例の処理は、

1. 2つ以上のバンドで撮られている領域のリストアップ
2. それぞれのデータの重なった領域の切り出し
3. それぞれのデータでの天体検出
4. 検出天体リストでの星、銀河の分類
5. 各領域での天体リストのクロスマッチ
6. 天体リストから star count へ

といった手順で分解できる。この処理は大きく分けてデータベース機能と画像処理機能の2つに大別できる。VO は地理的に分散した天文データベース/データアーカイブを対象にするので、分散を考慮しなければならないのはデータベース機能である。中でも複数のデータベースにまたがる検索が本質的である。そこで、VO の分散データベースへの問い合わせ言語として JVO Query Language を定義し、RDB (Relational Data Base) を用いて Subaru と 2 MASS のデータベースを構築し、その有効性を実証することを第1の目的にした。

天文データベースは世界中に分散しているため、インターネット接続された計算機間での連携を実現するミドルウェアが必要となる。世界の VO が連携するためにミドルウェアとして標準的な技術を使うことが重要である。現在ミドルウェアとして有望なのは唯一 GRID であり、世界の VO プロジェクトでも GRID が採用されている。GRID は Global Grid Forum (http://www.globalgridforum.org/L_About/about.htm 参照) で規格が制定されているが、昨年2月に GRID に Web service の機能を加える大きな拡張が提案された。プロトタイプには GRID 環境を実現するために Globus Toolkit version 2 を採用し、GRID が実際に機能することを確認し機能を評価することを第2の目的とした。

[裏へ続く]

```

create [materialized] view mytable as
  select c1.column1, c1.column2, c2.column1, c2.column2, ...,
         d1.BOX(POINT(c1.ra, c1.dec), width1, height1) as BlueImage,...
         c1.column3 / c2.column5 as flux_ratio, c3.*,...
  from catalog1 c1, catalog2 c2, catalog3 c3, ..., data d1,...
  where XMATCH(c1, c2, !c3...) < 3 arcsec [NEAREST | BRIGHTEST | ALL]
         and (c1.column1 - c2.column1) < 6.0 mag
         and BOX(POINT(ra0, dec0), width0, height0)
         and ...

```

図 1. JVO-QL コマンドの例 (1)

2.2 JVO Query Language

VO では様々な望遠鏡や観測装置で取得された観測データがアーカイブに保存され、データベース管理されていることを想定している。実際に使用されているデータベースシステム (DBMS) は RDB が一般的である。どのようなデータベースシステムが VO として必要かは慎重に検討する必要がある。画像データを扱える OODB (オブジェクト指向データベース) も候補であるが、先に述べた use case を分析した範囲では RDB で対応可能であると判断し、JVO プロトタイプでは RDB を内部システムとして採用することにした。

JVO では単一のインターフェイスで複数のデータサーバからカタログデータ、画像データを取得することができる。JVO システムに対するデータ請求要求を記述するインターフェイス言語となる JVO Query Language の文法として SQL (Structured Query Language) を拡張した。

図 1 にカタログデータを利用してそれらに対応する画像データを要求する検索の JVO-QL コマンドを示す。説明の詳細は省くが、主な拡張は where 節にある。XMATCH 演算子は複数のカタログをクロスマッチすることを意味し、

`XMATCH(c1, c2, !c3...) < 3 arcsec [NEAREST | BRIGHTEST | ALL]` は `c1` と `c2` を許容誤差 3 秒角の positional matching を行い、さらにそれらのうち `c3` には含まれていないものを選択することを意味する。

また、`(c1.column1 - c2.column1) < 6.0 mag` は通常の SQL 文と同様の条件式であるが、カタログのメタデータで指定されている単位に基づいて異なるカタログ間での単位変換を行った後に演算、比較を行う。

簡単な例として、図 2 に 2 色のデータ (`data1` と `data2`) があってその観測領域が重なっている部分のデータを取得する JVO-QL コマンドを示す。

```

select X.a, Y.a
  from data1.wavelength1 X,
         data2.wavelength2 Y
  where (X.AREA() OVERLAP Y.AREA()) as a

```

図 2. JVO-QL コマンドの例 (2)

ここで、`X.AREA() = data1.wavelength1.AREA()` は画像データアーカイブ `data1` の波長 `wavelength1` のデータがカバーしている領域を表す。

検索結果は International Virtual Observatory のために開発されたカタログデータの標準フォーマットである VOTable (<http://cdsweb.u-strasbg.fr/doc/VOTable/> 参照) で与えられる。

2.3 システム構成

プロトタイプでは Solaris OS の WS と Linux OS の PC という異機種のハードウェアを組み合わせたテスト環境を用い、DBMS は Oracle と PostgreSQL の 2 種類を使った。その他はフリーソフトウェアを使用した。GRID の実装としては Globus toolkit v2 に含まれている OpenSSL (Secure Socket Library)、OpenLDAP (LDAP Server)、wu-ftpd (ftp server) を使用した。Web サービスの機能は Globus toolkit v2 に含まれていないため SOAP (Simple Object Access Protocol) を採用した。

図 3 に JVO プロトタイプのシステム構成を示す。利用者は自分の好みの Web Browser で JVO Portal サーバにアクセスする。利用者認証後、JVO のメイン画面が現れる。この画面で JVO-QL で記述した仮想観測命令を JVO コントローラに送る。コントローラは命令を解析し、UDDI Registry を参照して要求されたサービスを提供するサーバを引き当て、各サーバの GSDL (WSDL の GRID 版: サービスのインターフェイスが定義されている) を参照して、仮想観測命令の実行手順を作成する。この手順に従い、GRID 経由で各サーバに実行要求を送る。予定されたサーバが何らかの原因で実行に失敗すると、コントローラは同じサービスを提供する別のサーバを探し、それに実行要求を出し直す。これを繰り返し最終的な結果が JVO サーバに出来上がる。

このプロトタイプでは DB サーバとして、2 MASS のカタログサーバと Subaru-SXDS のカタログ・画像サーバの 2 種類を実装した。一般に、天文 DB サーバは読み取り専用でデータ検索以外のサービスを期待できない。そのため、カタログの相互比較機能は JVO サーバの一つとして実装した。

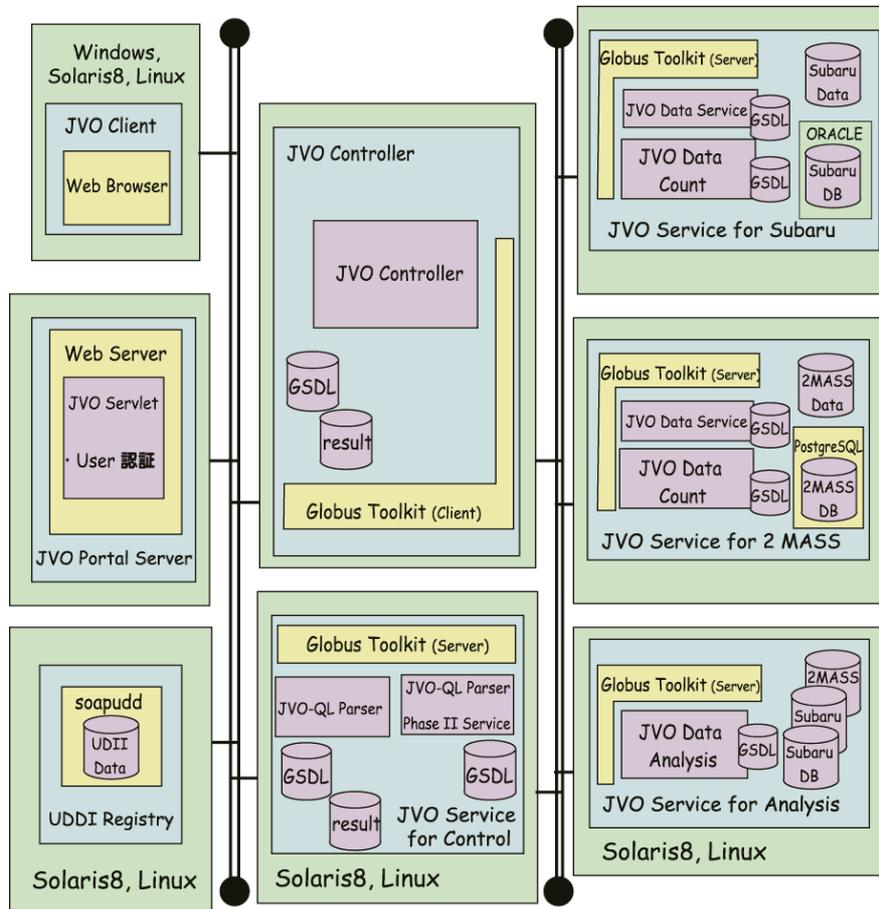


図3 JVO プロトタイプのシステム構成図

2.4 プロトタイプの評価

分散したデータベースをJVO-QLにより統一的に扱い、複数の独立なDB上のカタログデータの相互比較を行い、一致した天体の部分画像の取得ができることをJVOプロトタイプによって実証した。またソフトウェア技術の観点からは、GRID準拠の分散処理システムとして動作することを実証した。ただし、プロトタイプの実装開始時点で利用可能だったGRID基盤のGlobus Toolkit v2にはWebサービス機能がないため、その機能はSOAPで実現した。

大規模カタログの検索やカタログ間の相互比較にかかる処理時間は数秒で終わるものから何時間もかかるものまで予想がつかない。利用者はJVOに対する要求を出したら処理結果を待たずに操作を終了し、必要なときに処理経過をJVOに問い合わせるという使い方が望ましい。JVOプロトタイプではWebブラウザを用いてインターネットの世界からGRIDのPortalサーバを通してGRIDの世界のJVOシステムを使うという方式によって、利用者の操作部分とJVOの実際の仕事をするサーバ部分の分離を実現した。

実行性能はGRIDを使用しないで既存のサーバクライアント技術でシステムを構築したほうが格段に良いと思われる。プロトタイプではサーバに実行要求を出してから実際に起動されるまで数秒かかる。この性能の悪さはGlobus Toolkitに原因がある。

GRIDサービスの起動時の認証等のオーバーヘッドが一因となっており、起動時の認証を毎回行わず、キャッシュを利用するなどの改善策も知られている。プロトタイプによるGRIDの評価としては、実行性能はまだ不十分であるが、機能として使えないものではない。

3. JVOの今後

今回実装したのは分散したデータベースへの検索要求と、カタログ間の相互比較(クロスマッチ)機能だけで、そのほかのデータ解析機能は未実装である。実際のVOにはデータ解析機能が必須であり、これをどのようにJVOシステムに組み込むまたは、連携させるかが今後の大きな課題である。またプロトタイプには、セキュリティやデータ管理・エラー処理などといった機能は実装されていない。GRIDの観点から見ると、実用システムに必須なこれらの機能をGRIDによってどれだけ容易に実装できるかも課題である。

この初めてのプロトタイプの開発は富士通との共同研究として行った。JVO-QL関連の部分で国立天文台、GRIDサービスの枠組み部分を富士通という役割分担で開発した。GRIDという新しいネットワーク分散コンピューティング技術を基盤にした枠組みの上に、VOという数値宇宙の観測装置を構築するのがJVOプロジェクトである。データベース天文学の手法の確立とその道具となるシステムの開発という2つのテーマは相互依存性が強く、同時に研究開発を進めることが重要である。データベース天文学推進室の数名のメンバーだけではできることが限られてしまうが、ソフトウェア開発を専門とするメーカーとの共同研究は非常に有効であった。

VOの真髄は波長を越えた天文学にある。また、JVOはGRIDという次世代技術を採用したチャレンジングなプロジェクトである。電波からガンマ線まで広い範囲の天文研究者がVOに興味を持ちJVOの開発に協力しようと手を上げること、更に、ソフトウェア開発の面からJVOに協力しようという方が読者の中に現れることを期待してこの紹介を終わる。

(国立天文台 データベース天文学推進室
水本 好彦)

新メンバー自己紹介

2003年2月1日よりPLAINセンターの助手に就任しました田村隆幸です。

PLAINセンターでは、主にDARTSを担当する予定です。宇宙研の衛星のデータから研究者が最大限の成果を上げることを目標に、DARTSを管理・発展させていきたいと思っております。今後打ち上げられるAstro-FやAstro-EIIのデータが、宇宙研の内外に関わらず、ストレスなく解析できるようなシステムを提供することが最初の仕事になると思っております。このためには、データ利用者の声が不可欠であります。DARTSに関して、質問、希望などありましたら、ぜひお寄せ下さい(comments@darts.isas.ac.jp)。

さて、私はPLAINセンターに来る前には、オランダのSRON (Space Research Organization Netherlands) でポスドクを3年間やっておりました。ちょうど私が、SRONに赴任する数カ月前に、

ヨーロッパのX線衛星・Newtonが打ち上がりました。私は主に検出器(X線分光器;RGS)の軌道上での較正と銀河団の観測を行ないました。SRONに行く前は、宇宙研のX線グループでAstro-Eの打ち上げ直前まで、その試験に参加しました。その前は、東大の物理・牧島研にて、Astro-EのX線検出器(HXD)の開発と「あすか」衛星による銀河団の研究を行なっておりました。というわけで、宇宙研のX線グループには知っている方(特に男性)が多いので、今後はできるだけそれ以外の方との交流を深めていきたいと思っております(笑)。気軽に声をかけて頂ければ、とても嬉しいです。前に宇宙研に居た時には、昼休みに中庭でバレーボールをやっていましたが、またやりませんか？

それでは皆さん、楽しくいきましょう。

(田村 隆幸)

大型計算機に関するお知らせ

1. 課題更新手続きについて

現在利用されている課題の有効期限は、3月31日までとなっておりますので、次年度も引き続き利用される課題の更新手続きは、3月14日(金)までをお願いします。課題更新用紙は、2月21日(金)にお送りしましたのでご利用ください。

* 課題更新用申請用紙のフォーマットについて *
課題更新用紙は3種類ありますので、申請目的にあった用紙をご利用ください。

- (1) GS8300/10N MSP用 新規申請用紙
 - (2) UNIX用 新規申請用紙
 - (3) 課題更新MSP・UNIX 更新用申請用紙
- ※MSPとUNIXは、各々別の管理をしていますので、申請時には各々予算登録をしてください。
※更新用紙には同一予算費目、同一内容のMSP・UNIXのセット課題を上段・下段に記入し、提出してください。

※新規課題申請用紙は、A棟4F、7F、B棟1F、C棟2F、D棟2Fの各サブステーションのレターケースの中にあります。

2. 大型計算機年度末処理について

毎年行われている年度末処理を、今年は3月30日(日)・31日(月)の2日間に行います。年度末処理では、次の事項を行います。

- ・システムバックアップ
 - ・年間ジョブ集計、課金集計処理
 - ・課題登録、更新処理
 - ・マスタファイルの切り替え
 - ・CEファーム改版作業
 - ・重大障害修正作業と確認
- ※VPPは29日(土)午前3時にジョブキュー停止になります。
※VPP以外は特に事前の停止は行いませんので、各自の判断で入力してください。
3月30日(日)の9:00までに終了しないジョブはキャンセルします。
※GS8300/10NのMSPにおけるジョブ出力結果は、3月28日(金)までに取り出してください。また、取り出しできなかったジョブ出力は、MTにバック

アップを取ってありますので、必要な方は4月1日以後、内線8391高橋、林まで連絡してください。
※課題の更新をしない方は、3月28日(金)までにプログラム・データのMT吸い上げを行い、ファイルの消去をしてください。MSPの場合はデータのバックアップユーティリティにはSASAR、SASMVがあります。

3. 3月・4月の計算機年度末処理及び保守作業予定

3月・4月は、下記の通り年度末処理及び保守作業を予定しております。

計算機名	3/29(土) 18:00~	3/30(日)~3/31(月) 8:00~24:00
GS 8300/10 N		M 年度末処理
VPP800/12	M 年度末処理	M 年度末処理
その他の Server		M 年度末処理

M: システムメンテナンス

○VPP800は3月28日(金)からSIMPLEXモードに移行します。

- ジョブキュー停止は29日(土)午前3時です。
- 3月に予定されていたVPP800の保守作業及び4月に予定されていたAlphaサーバの保守作業は年度末保守が予定されているので中止します。

4. 大型計算機のリプレースについて

今年夏の計算機リプレースに、以下の計算機・端末等が含まれています。これらの装置上で使われているプログラム等は、リプレースまでに他のプラットホームへの移植作業等が必要となります。

- ・センター側MSP計算機(GS8300): 撤去となります。後継機は予定されていません。
- ・高機能端末(FMV): 上記MSP計算機の撤去に伴い、高機能端末も全て撤去となります。買い取りは出来ません。(本体、ディスプレイ等に「fmv99」で始まるラベルが貼られた装置が該当します)
- ・Alphaサーバ: 後継機は他のプラットホームに変更予定です。

5. 大型計算機の相談窓口について

大型計算機利用上の質問・トラブルなどは高橋氏・林氏(内線8391)、ネットワーク関係の質問・トラブルなどはPLAINセンター本田秀之(RN1261・内線8073)までお願いします。(三浦 昭)